

UNIVERSIDADE FEDERAL DO PARANÁ

DIRCEU NUNES DE FARIA JUNIOR

AUDITORIA EM BIG DATA

CURITIBA
2020

DIRCEU NUNES DE FARIA JUNIOR

AUDITORIA EM BIG DATA

Monografia apresentada ao Departamento de Ciências Contábeis, do Setor de Ciências Sociais Aplicadas, da Universidade Federal do Paraná, como pré-requisito para obtenção do título de Especialista em MBA – Auditoria Integral.

Orientadora: Prof^a. Dr^a. Denise Fukumi Tsunoda

CURITIBA
2020

AGRADECIMENTOS

Quero agradecer em primeiro lugar a Deus, por ter me guiado durante esta etapa e colocado pessoas essenciais em minha vida.

À minha esposa, que acreditou em minha capacidade e nunca me deixou desistir. Por muitas vezes tornou-se mãe e pai de nosso pequeno, tentando suprir minha ausência. Mas, sempre com um sorriso no rosto, me recebeu em nossa casa.

Ao meu filho, que serviu de inspiração para que eu retornasse à sala de aula. Mesmo pequeninho, entendeu que o papai precisava estudar.

À minha família, que sempre me motivou, entendeu as minhas faltas e momentos de afastamento e reclusão.

Aos amigos que muito estudaram comigo durante vários sábados, foram momentos de imensa alegria e certamente vocês fazem parte dessa história.

À minha orientadora, Profa. Dra. Denise Tsunoda, por todos os momentos de paciência, compreensão e competência. Sem você, com certeza essa monografia não existiria.

Agradeço ainda a todos os professores do Curso de MBA em Auditoria Integral da Universidade Federal do Paraná em Curitiba, por terem dividido o seu conhecimento e pela paciência com nossa turma. Com certeza, tornaram-se nossos amigos também.

Finalmente, agradeço a todas as pessoas que de alguma forma contribuíram para a conclusão deste MBA.

Muito Obrigado.

RESUMO

Este estudo propõe mapear as principais tendências, riscos e benefícios na utilização de *Big Data* nos processos de auditoria, ao mesmo tempo em que descreve e delimita estes conceitos na sua utilização contemporânea. A pesquisa enfatiza o cenário brasileiro, mas traz dados do cenário internacional para ajudar a contextualizar o que ocorre no Brasil dentro do tema. O método utilizado é a revisão de literatura indexada em bases de dados acadêmicos, incorporando também referências derivadas de pesquisa livre. As referências selecionadas têm seus principais conceitos apresentados na fundamentação, seguidos por uma discussão e articulação crítica entre os dados encontrados e as práticas de auditoria atual. Os resultados apontam que a auditoria digital com o uso de *Big Data* ainda se encontra em estágio de desenvolvimento, tanto da tecnologia quanto dos recursos humanos para empregá-la. Esta observação é válida para o cenário brasileiro e também para o que ocorre em nível internacional, sendo que o Brasil possui maiores complicações no que concerne ao dispêndio financeiro com novas tecnologias e formação continuada nas pequenas e médias empresas. Porém, observa-se um estímulo ao desenvolvimento da auditoria digital no contexto nacional através de algumas iniciativas de *e-government*, como a obrigatoriedade do uso do Sistema Público de Escrituração Digital – SPED. Também se nota que a participação do auditor humano não deve ser dispensada em um futuro próximo pois, mesmo com o rápido avanço das tecnologias envolvidas, o auditor assume progressivamente tarefas que envolvem decisões éticas e morais que não podem ser realizadas adequadamente pela Inteligência Artificial. A sugestão para pesquisa futuras é que seja ampliado o escopo de avaliação da implementação da auditoria digital, principalmente com estudos de caso que identifiquem problemas práticos da aplicação da IA à análise de *Big Data*.

Palavras-chave: Aprendizagem de Máquina. Auditoria. *Big Data*. Inteligência Artificial.

ABSTRACT

This study proposes to map the main trends, risks and benefits in the use of Big Data in the auditing process, while also describing these concepts in their contemporary use. The research emphasizes the Brazilian scenario, but brings data from the international practices to help contextualize what occurs in Brazil within the theme. The method used is the literature review of indexed articles in academic databases, also incorporating references derived from free research. The selected references have their main concepts presented in the rationale, followed by a discussion and critical articulation between the data found and the current audit practices. The results show that digital auditing with the use of Big Data is still in the development stage, both of technology and human resources to employ it. This observation is valid for the Brazilian scenario and also for what occurs at an international level, with Brazil presenting more difficulties regarding financial expenditure on new technologies and employee education in small and medium-sized companies. However, there is a stimulus to the development of digital auditing in the national context through some e-government initiatives, such as the mandatory use of the Public Digital Bookkeeping System (Sistema Público de Escrituração Digital – SPED). It is also noted that the participation of the human auditor should not be dispensed with in the near future because, even with the rapid advancement of the technologies involved, the auditor progressively takes on tasks that involve ethical and moral decisions that cannot be carried out properly by Artificial Intelligence. The suggestion for future research is to expand the scope of evaluation of the implementation of digital auditing, mainly with case studies that identify practical problems in the application of AI to Big Data analysis.

Keywords: Artificial intelligence. Audit. Big Data. Machine Learning.

LISTA DE ILUSTRAÇÕES

QUADRO 1 – DEFINIÇÕES PARA BIG DATA.....	13
FIGURA 1 – TEMAS RELACIONADOS A <i>BIG DATA</i> NA LITERATURA ACADÊMICA.....	04
FIGURA 2 – PROCESSOS DE <i>BIG DATA</i>	20
TABELA 1 – TECNOLOGIAS APLICADAS ÀS ATIVIDADES CONTÁBEIS.....	26
QUADRO 2 – TECNOLOGIAS REFERENTES AO USO DE INTELIGÊNCIA ARTIFICIAL.....	27
QUADRO 3 – COMPARAÇÃO ENTRE O PROCESSO DE AUDITORIA TRADICIONAL E AUDITORIA COM INTELIGÊNCIA ARTIFICIAL (FASE INICIAL).....	28
QUADRO 4 – COMPARAÇÃO ENTRE O PROCESSO DE AUDITORIA TRADICIONAL E AUDITORIA COM INTELIGÊNCIA ARTIFICIAL (FASE EXECUÇÃO).....	29
QUADRO 5 – COMPARAÇÃO ENTRE O PROCESSO DE AUDITORIA TRADICIONAL E AUDITORIA COM INTELIGÊNCIA ARTIFICIAL (FASE FINAL).....	31
QUADRO 6 – PROCESSO DE BUSCA DAS REFERÊNCIAS.....	34
QUADRO 7 – PROCESSO DE SELEÇÃO DAS REFERÊNCIAS.....	35
QUADRO 8 – LISTA DE REFERÊNCIAS SELECIONADAS E EMPREGADAS.....	36

SUMÁRIO

1	INTRODUÇÃO.....	07
1.1	PROBLEMA E JUSTIFICATIVA.....	07
1.2	OBJETIVOS.....	09
1.3	ESTRUTURA DA MONOGRAFIA.....	09
2	FUNDAMENTAÇÃO TEÓRICA.....	11
2.1	<i>BIG DATA</i>	11
2.1.1	<i>Big Data</i> nas Organizações.....	12
2.1.2	<i>Big Data</i> e Inteligência Organizacional.....	16
2.2	AUDITORIA DIGITAL.....	19
2.2.1	Auditoria Digital em <i>Big Data</i>	20
2.2.2	Riscos e Possíveis Benefícios.....	22
2.2.3	Principais Instrumentos e Processos.....	25
3	PROCEDIMENTOS METODOLÓGICOS.....	33
3.1	METODOLOGIA.....	33
3.2	MATERIAIS.....	34
3.3	MÉTODOS DE ANÁLISE.....	36
4	DISCUSSÃO.....	38
5	CONSIDERAÇÕES FINAIS.....	44
	REFERÊNCIAS.....	46

1 INTRODUÇÃO

Os sistemas informatizados de gerenciamento administrativo não são novos, e registra-se que desde a década de 60 a tecnologia da informação e comunicação já ensaiava sua presença nas organizações, sempre visando à maior velocidade e confiabilidade nos processos realizados. As áreas de contabilidade e controladoria, devido à sua importância vital para qualquer empresa, sempre foram depositários de muitas horas de trabalho pelos funcionários internos e externos, engajados em manter regularizada a documentação fiscal e contábil, e garantindo os registros confiáveis das movimentações dos recursos organizacionais. A grande quantidade de trabalho envolvida nestes processos tornou a área um alvo natural para a tentativa de automatização e, posteriormente, informatização. (ALVES; FREITAS, 2017).

As Técnicas de Auditoria Assistidas por Computador (TAAC) já acompanham os trabalhos do auditor há alguns anos, mas sua tendência de uso passou por grande incremento devido à integração com o *Big Data*, um novo universo de dados em grandes volumes, consequente do avanço de tecnologia de coleta de dados e também da rede mundial de computadores. A integração dos sistemas informatizados com essa fonte de dados gerou um novo cenário para os escritórios de auditoria externa, auditores internos e controladores, que se encontram ainda em fase de adaptação ao emprego destas novas tecnologias.

De acordo com as mudanças observadas nas últimas décadas – e principalmente ao caráter de aceleração verificado nas mudanças das auditorias nos últimos anos – este trabalho propõe apresentar um estudo sobre como o *Big Data* está se vinculando aos processos de auditoria tradicionais, criando novas práticas, oferecendo novas modalidades de informação, modificando seu caráter de pontual para contínuo e exigindo novas habilidades e competências do auditor.

1.1 PROBLEMA E JUSTIFICATIVA

Costa e Dutra (2014) apontam que a auditoria financeira é usualmente aplicada em áreas de alto risco, utilizando amostragem estatística que permite a realização dos

procedimentos com maiores benefícios e custos reduzidos. Além disso, também é comum que a auditoria acompanhe o calendário anual de demonstrações financeiras, apresentando uma periodicidade que coloca a empresa em auditoria por certos períodos, mas que exhibe relaxamento destes controles na maior parte do tempo.

Entretanto, o cenário do sistema financeiro mundial mudou rapidamente nas duas últimas décadas, com informações disponibilizadas de forma mais rápida e em grande volume. Independentemente desta mudança, as informações continuam necessitando de validação, para que sejam utilizadas em escolhas racionais por investidores, participantes do mercado, departamentos financeiros e gestão corporativa. O cenário impõe uma nova forma de auditoria, que se aproxima gradativamente do conceito de transparência e prestação de contas em tempo real, com controle interno e externo constante. (COSTA; DUTRA, 2014).

Deve-se considerar ainda que o grande volume de dados coloca novas responsabilidades para o profissional auditor, e o cruzamento das informações do tipo *Big Data* com o Controle Interno corporativo traz desafios, riscos e muda o âmbito de análise de dados. (VIEIRA, 2016). É possível ocorrer também o envolvimento de novos setores e departamentos da empresa na análise dos dados, dando visibilidade a fragilidades e potencialidades que só emergem quando se considera um grande volume de informações, tradicionalmente ignoradas ou deixadas de lado na amostragem. Estes motivos indicam que é necessária reflexão e estratégia anterior à instalação do processo de auditoria em *Big Data*.

O estudo de Pedrosa, Laureano e Costa (2015) aponta que os principais motivadores para o uso de ferramentas digitais entre os auditores estão relacionados ao seu nível de formação profissional e ao fato de trabalharem em empresas de maior porte. Os autores concluem que é importante aumentar a oportunidade de formação dos auditores – principalmente os mais jovens – para que tenham segurança a respeito da aplicação e uso das ferramentas digitais de auditoria.

Sendo assim, este trabalho procura expandir os estudos sobre a auditoria digital praticada atualmente, identificando seus riscos, benefícios, potencial de expansão e ferramentas aplicadas na realidade nacional. Esta modalidade de estudo se mostra necessária devido à expectativa de crescimento do número de dados financeiros gerados

por empresas de todos os portes, e ao próprio cenário de regulamentação fiscal, que caminha para demandar as informações financeiras de forma contínua, assim como já é exigido por gestores e grandes investidores. Considerando todas as questões colocadas a pergunta de pesquisa é: como o Brasil se posiciona diante do cenário internacional de ferramentas e processos de auditoria digital?

1.2 OBJETIVOS

Este trabalho apresenta como objetivo geral discutir processos e ferramentas de auditoria digital que se aplicam ao cenário brasileiro, considerando a utilização de Big Data como insumo fundamental do controle interno e externo. Além disso, também propõe os seguintes objetivos específicos:

- descrever e delimitar o conceito de *Big Data*;
- compreender o papel do *Big Data* nas corporações, especialmente no cenário brasileiro;
- descrever e delimitar o conceito de auditoria digital utilizando *Big Data*;
- mapear o potencial e os riscos de utilização de *Big Data* em auditoria digital pontual e contínua nas corporações.

1.3 ESTRUTURA DA MONOGRAFIA

A estrutura deste trabalho é dividida em cinco segmentos principais, dos quais faz parte esta Introdução, ou Capítulo 1. O próximo segmento se refere à Fundamentação Teórica (Capítulo 2), no qual são apresentados e detalhados os conceitos de auditoria digital e *Big Data*, trazendo também algumas ferramentas e procesos representativos da tendência atual de realização da auditoria contínua e com grandes volumes de dados. Em seguida, os Procedimentos Metodológicos compõem o Capítulo 3, que descreve os passos de pesquisa de referências, seleção dos dados e análise dos resultados empregados nesta monografia. O Capítulo 4, Discussão, integra os conceitos apresentados da Fundamentação, trazendo estudos sobre a aplicação da

auditoria digital e o emprego de Big Data em cenários e contextos contemporâneos. A monografia é finalizada com a Considerações Finais (Capítulo 5) e a lista de Referências.

2 FUNDAMENTAÇÃO TEÓRICA

Esta seção apresenta os principais conceitos relacionados à pesquisa em pauta, dentre os quais destacam-se: big data e auditoria digital.

2.1 *BIG DATA*

De acordo com De Mauro, Greco e Grimaldi (2016), o termo "*Big Data*", ocupa cada vez mais espaço em artigos e relatórios referentes às práticas e pesquisas na área de Tecnologia da Informação. Devido à presença crescente das tecnologias digitais e das aplicações baseadas em coleta e análise de dados, também se observa que o termo se disseminou entre disciplinas como sociologia, biologia, economia, medicina, administração e gestão. Entretanto, apesar do aumento do uso dos conceitos relacionados ao tratamento de grandes quantidades de dados, o vocabulário pertinente a este tema não foi unificado para o uso em tantas áreas e situações diferentes.

A explicação mais simples para *Big Data* se refere ao processamento de um volume muito grande de dados estruturados ou não estruturados (ANDREU-PEREZ et al., 2015); de fato, um volume tão grande que precisa de sistemas informatizados para que possa ser transformado em informação útil. Entretanto, as definições mais detalhadas abrangem também outros "Vs", além do Volume.

Embora existam divergências quanto ao número de características que desenham o que é ou não *Big Data*, este trabalho opta por 6 Vs básicos: 1) *volume*, já citado; 2) *valor*, o processamento dos dados massivos deve gerar informação relevante e empregável; 3) *velocidade*, os dados a serem avaliados são gerados de maneira muito rápida, e precisam também de análise ágil ou em tempo real; 4) *variedade*, os dados são heterogêneos em sua natureza, e são provenientes de diferentes fontes; 5) *veracidade*, as fontes de informação e os dados devem ser confiáveis; e 6) *variabilidade*, diferente da variedade (que se refere à natureza dos dados), a Variabilidade diz respeito aos modelos de análise, que podem utilizar modelos estatísticos tradicionais, mas que também precisam de abordagem não-determinística

para complementar as análises e gerar novas categorias de informação. (ANDREU-PEREZ et al., 2015).

2.1.1 *Big Data* nas Organizações

A criação e expansão do *Big Data* foi motivada inicialmente pelo crescimento rápido dos cenários nos quais dados eram criados, compartilhados e utilizados. A década de 1990 foi o início do movimento de *digitalização*, com a conversão de sinais analógicos em sinais digitais, quando surgiram as primeiras ferramentas comerciais de reconhecimento ótico de caracteres (OCR). Estes softwares permitiram que projetos de digitalização em massa se tornassem realidade, transformando bibliotecas tradicionais analógicas (impressas) em arquivos digitais acessíveis e legíveis por máquinas. Com a conversão de caracteres analógicos para o formato digital tornou-se possível organizá-los em conjuntos de dados mais estruturados, oferecendo um inédito ponto de vista em nível macro, mais abrangente do que a leitura sequencial de textos impressos, revelando novas tendências e padrões cuja visualização seria impossível se todos os dados permanecessem em formato analógico. (De MAURO, GRECO E GRIMALDI, 2016).

Considera-se que o *Big Data* atinge o ápice de seu desenvolvimento com o crescimento das mídias sociais e weblogs, alocando atividades básicas de análise e inteligência de negócios (*Business Intelligence*) sobre novas fontes de dados (como aplicativos de entregas por celular ou sites de relacionamento), com análises profundas e em tempo real das informações disponibilizadas pelas transações e pelos usuários. Entretanto, o volume de dados gerados no mundo digital cresceu exponencialmente e tornou-se difícil gerenciá-lo usando as tecnologias de armazenamento de dados que estavam disponíveis. Este cenário criou condições para o que se convencionou chamar de *Big Data*, que apresenta definições diversas em diferentes estudos como indicado no Quadro 1:

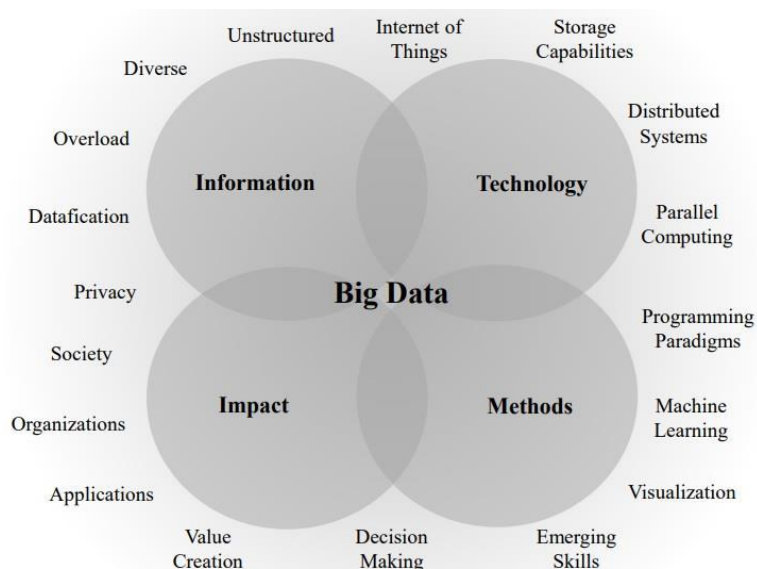
QUADRO 1 – DEFINIÇÕES PARA BIG DATA

Autores	Definições
Kim, Trimi, e Ji-Hyong (2014)	Trata-se de um termo geral para a enorme quantidade de dados digitais coletados a partir de todo os tipos de fontes.
Mahrt e Scharkow (2013)	Denotam um maior conjunto de dados ao longo do tempo, conjuntos de dados estes que são grandes demais para serem manipulados por infraestruturas de armazenamento e processamento regulares.
Davenport (2012) e Kwon (2014)	Dados demasiadamente volumosos ou muito desestruturados para serem gerenciados e analisados através de meios tradicionais
Di Martino et al. (2014)	Refere-se ao conjunto de dados cujo tamanho está além da habilidade de ferramentas típicas de banco de dados em capturar, gerenciar e analisar.
Rajesh (2013)	São conjuntos de dados que são tão grandes que se tornam difíceis de trabalhar com o uso de ferramentas atualmente disponíveis.
Mahrt e Scharkow (2013)	Denotam um maior conjunto de dados ao longo do tempo, conjuntos de dados estes que são grandes demais para serem manipulados por infraestruturas de armazenamento e processamento regulares.

FONTE: FREITAS JUNIOR et al. (2016, p. 532).

Como pode ser observado, embora sejam diferentes as definições convergem para a consideração do conceito como sempre agregando uma grande quantidade de dados a serem analisados, e que geram desafios devido ao volume. Este fato é rotineiramente testemunhado nas empresas que trabalham com dados em tempo real, e capturam trilhões de bytes de informações sobre seus clientes, fornecedores e operações. Além disso, existe atualmente uma quantidade enorme de sensores conectados em rede que fazem parte de dispositivos como telefones celulares e automóveis, detectando, criando e comunicando dados continuamente. (AJAH; NWEKE, 2019).

De Mauro, Greco e Grimaldi (2016) apontam que, conforme revisão de literatura, outros temas também estão relacionados aos estudos em Big Data, como demonstrado na Figura 1:

FIGURA 1 – TEMAS RELACIONADOS A *BIG DATA* NA LITERATURA ACADÊMICA

FONTE: De MAURO, GRECO E GRIMALDI (2016, p. 126).

Como ilustrado pelos autores, são quatro grandes áreas de pesquisas relacionadas a *Big Data*: Informação, Tecnologia, Métodos e Impacto. Em cada uma destas diferentes temas são abordados, sendo que estes não são mutuamente excludentes, como mostrado pela intersecção dos grupos. Assim, por exemplo, as questões referentes à privacidade dos usuários de novas tecnologias referem-se tanto à área de geração e manejo da informação quanto à área que trata do impacto de seu uso e implementação. Já na área de tecnologia as pesquisas podem ir de temas bastante específicos, como o desenvolvimento de novas formas de armazenamento com o fim de aumentar a capacidade, até à novíssima Internet das Coisas (*Internet of Things - IoT*).

O uso generalizado dos smartphones e de sites e aplicativos de redes sociais continuarão alimentando o crescimento deste tipo de informação que forma o *Big Data*: grandes agrupamentos de dados que podem ser capturados, comunicados com velocidade, sintetizados em blocos ou categorias, armazenados e analisados. Para as empresas, esta quantidade de dados a ser absorvida nos processos produtivos provoca necessidades e preocupações:

A expressão “*Big Data*” não trata apenas de designar grandes volumes de dados. Quando uma organização implementa uma solução *Big Data* significa que adotou mecanismos para lidar, analisar e retirar mais-valias do facto de haver tanta informação. O tratamento desta informação exige a implementação de circuitos, a criação de interfaces, a monitorização de processos. Ou, de um modo mais simples, exige preparação, trabalho e estratégia. Tradicionalmente para alguns estudiosos a tecnologia existe para assistir precisamente nessas áreas. É muitas vezes no sistema de informação que está a diferença entre analisar correta ou incorretamente a informação. (VIEIRA, 2016, p. 7).

Considera-se que o *Big Data* faz parte de todos os setores e funções da economia global, adquirindo importância próxima à dos insumos a serem beneficiados na indústria, ou à do capital humano no setor de serviços. Uma parcela substancial da atividade econômica moderna, da inovação e do crescimento não poderiam ocorrer sem este tipo de agregamento de dados e sua análise efetiva. (MANYIKA et al., 2011).

Diferentes estudos (BRYNJOLFSSON; HITT; KIM, 2011; MANYIKA et al., 2011) já relatavam no início da década de 2010 que a enorme quantidade de dados brutos gerados precisava de tecnologia de *Big Data* para análise. Naquela época já existiam 5 bilhões de aparelhos celulares em uso no mundo, e estima-se que cada um deles cria um terabyte de dados anualmente. Somente o *Facebook* já estimava 30 bilhões de informações compartilhadas por mês entre seus usuários. Quando o olhar é voltado para as linhas de negócio mais tradicionais observam-se dados como os do Wal-Mart, que processa mais de um milhão de transações de clientes por hora e armazena 2,5 petabytes de dados de clientes.

Estes dados cresceram desde o início da década, em todos os setores econômicos, alimentados por mudanças tecnológicas e atividades cada vez mais intensas no comércio eletrônico (*e-commerce*). Ao mesmo tempo, novas tecnologias para a captação de dados – como a Internet das Coisas – e para a gestão e compartilhamento destes mesmos dados – como a computação em nuvem (*cloud computing*) – fornecem armazenamento persistente e transformação de dados em uma escala que até pouco tempo atrás não era possível. (SAIF; WAZIR, 2018).

2.1.2 *Big Data* e Inteligência Organizacional

As abordagens de *Big Data* e análise de negócios foram desenvolvidas e implementadas de forma relativamente recente, com o objetivo de analisar um grande volume de dados gerados por diferentes organizações comerciais e industriais:

A era do big data compreende diferentes facetas e processos, que envolvem procedimentos de geração, seleção e manipulação de dados. Distintos campos técnicos e do conhecimento são afetados por esse novo ambiente, o que gera o desenvolvimento de estudos e aplicações em distintas áreas científicas, como computação, elétrica, ciências sociais, dentre outras. Ademais, estudos em estratégia e gestão têm explorado o tema como um vetor revolucionário dos negócios e do modo como as empresas adquirem vantagens competitivas. (FURLAN; LAURINDO, 2017, p. 92).

Por este motivo, o cenário atual dita que toda empresa precisa de uma maneira mais ágil e eficiente para visualizar volumes crescentes de dados que mudam rapidamente. A análise de dados em tempo real ajuda as organizações a olhar de forma mais consistente para o passado e prever o futuro. Esta se configura como uma análise em 4 níveis, e é composta pela necessidade de saber o que ocorreu (nível descritivo), entender por que aconteceu (nível diagnóstico), olhar para o que ainda pode acontecer (nível preditivo) e, finalmente, influenciar nas possíveis ocorrências futuras (nível prescritivo). Certamente os quatro níveis de análise trazem enormes benefícios para a empresa que os utiliza de forma eficiente, mas são progressivamente mais difíceis de implementar e usar, sendo que o nível descritivo é o mais fácil, e o nível prescritivo o mais complexo. (AJAH; NWEKE, 2019).

Deve-se ainda considerar que o uso de *Big Data* não se conforma apenas como uma forma para obter eficiência nas operações comerciais ou melhores resultados econômicos. Também existem oportunidades importantes para a melhoria do padrão de vida dos trabalhadores e de diversos setores organizacionais da empresa. Isso pode incluir, por exemplo, a prestação de um melhor serviço de assistência médica, crescimento da média de educação, segurança adicional nas operações comerciais e garantia de boa governança corporativa. (AJAH; NWEKE, 2019). Para isto, coloca-se para as empresas a necessidade de reestruturação interna:

Quanto mais dados surgem, potencialmente mais informação e conhecimento se gera, mas mais difíceis se torna aceder e processar os dados. E, principalmente, de os usar para tomar as melhores decisões. É por isso que as organizações têm mesmo de estar preparadas para esta nova realidade. As que melhor trabalharem esta informação, ou seja, que produzirem mais rapidamente resumos e sínteses para tomar decisões, vão adiantar-se em relação às outras. Muitas vezes o que acontece é que a informação até já existe nos sistemas, mas como os humanos a desconhecem não puderam decidir sobre ela e perderam assim a oportunidade. Apenas os sistemas integrados têm a capacidade de identificar toda a informação útil e mostrá-la aos utilizadores. (VIEIRA, 2016, p. 8).

Assim, o uso de *Big Data* também apresenta importante potencial para ajudar os responsáveis pelo negócio com uma visão mais estruturada das políticas que podem oferecer um terreno seguro aos investidores.

De acordo com Grover et al. (2018), quando os gestores têm acesso às informações detalhadas sobre processos de negócios, provenientes da análise de um grande volume de dados, tais informações são subsídios que podem ser usados para melhorar sua eficiência, produtividade, acessibilidade e disponibilidade. Também se torna possível transformar os fluxos de informações dos processos de negócios, que podem ser aproveitados para criar vantagem competitiva e diminuir a concorrência. A análise de *Big Data* proporciona informações sobre produtos e serviços, analisando dados gerados internamente e o conteúdo externo gerado pelo usuário ou comprador:

Mais do que apenas se concentrar em uma medida específica de mercado, os sistemas de avaliação do desempenho de mercado devem preocupar-se em agregar combinações de diferentes tipos de métricas, com foco nas reações dos clientes e no impacto do produto no mercado juntamente com métricas financeiras e diferenciais de marketing, a fim de levar em consideração as perspectivas de lucro e de valor de longo prazo. (CASALINHO, 2015, p. 162).

Com tais informações, a empresa pode aumentar a diferenciação de produtos e serviços, ajustar o preço daquilo que é oferecido e desenvolver soluções inovadoras para os clientes. Da mesma forma, informações sobre clientes e mercados podem ser utilizadas no aumento da satisfação e promoção da fidelidade do cliente, atraindo novos interessados e envolvendo os fornecedores na criação de nichos de mercado. A análise de *Big Data* também pode trazer insights sobre o desempenho da organização como um todo, contribuindo para a inteligência organizacional e tomada de decisões; em tal contexto, aumentam as condições para desenvolvimento de uma estrutura

organizacional dinâmica, que responde de forma mais rápida às mudanças de mercado e demandas ambientais, o que leva a incrementos progressivos da utilização da capacidade e aumento do retorno sobre os ativos. (GROVER et al., 2018).

Ainda sobre o tema da competição organizacional baseada em *Big Data*, a pesquisa realizada por Gimenez (2018) sobre o tema, entre empresas brasileiras, foi segmentada em quatro objetivos específicos: 1) Avaliar as relações entre dimensões de *Big Data* com etapas do ciclo de Inteligência Competitiva; 2) Verificar se existe relação entre o ciclo de Inteligência Competitiva e fluxo de informações em *Big Data*; 3) Identificar resultados dos projetos consequentes do uso de *Big Data* em Inteligência Competitiva; e 4) Identificar tecnologias afetas ao crescimento de dados.

O autor relata que o uso das tecnologias para tratamento e emprego de *Big Data* de fato ocorre nas empresas brasileiras, identificando 13 ferramentas, sendo 8 delas citadas pelos entrevistados componentes das empresas abordadas. Com o uso das ferramentas, foi verificada a correspondência entre as dimensões de *Big Data* com as etapas da Inteligência Competitiva, e os resultados apresentados pelo emprego das ferramentas foram constatados como derivados da perspectiva de tratamento de dados em larga escala, o que seria impossível com uso das ferramentas tradicionais. Os ganhos específicos vão desde geração de conhecimento mais preciso, beneficiando a cadeia produtiva, até a previsão de falhas em equipamentos, como fator profilático que resulta em economia para a organização. (GIMENEZ, 2018).

De forma complementar, Esteban-Navarro e García-Madurga (2019) apresentam um estudo exploratório, a partir de entrevistas e grupos de discussão realizados com diretores e gestores de empresas espanholas. O estudo indica que a percepção da Inteligência Competitiva por estes executivos está estreitamente vinculada à análise de dados em grande escala, e que são necessárias mudanças que incluem até mesmo a própria cultura corporativa, já que todos os setores e departamentos da empresa precisam atribuir mais importância aos dados, para que estes se tornem informação e finalmente conhecimento. Porém, para que isto ocorra, também seria importante distribuir novas e melhores ferramentas de busca e organização de dados entre os vários departamentos, permitindo a formação de *clusters* de inteligência dentro da empresa que gerariam soluções para os problemas

que atravessam toda a cadeia de negócio. E, certamente, isto também requer formação de pessoa especializado para lidar com o novo fluxo de dados, gerando informações úteis a partir deles e com capacidade para analisá-los e transformá-los em conhecimento incorporado aos processos e produtos.

Embora as pesquisas demonstrem que o uso de *Big Data* é essencial para manter a competitividade das organizações, é importante notar que no cenário nacional este ainda se configura como um fator relativamente novo. O uso eficiente das informações ainda é um desafio crítico para grande parte dos sistemas de gestão, apresentando ainda dificuldades para integrar-se de forma eficiente às práticas de inovação nas empresas. Como indicado por Fusco e Mucheroni (2017), as organizações ainda precisam passar por uma evolução nos processos de gestão e construção contínua de novos conhecimentos, para que estes tenham efeitos positivos nos produtos e serviços; e, para atingir tal objetivo, novas investigações e pesquisas ainda são necessárias.

2.2 AUDITORIA DIGITAL

"Auditoria" não significa uma atividade ou conceito rígido, já que se observa que ele passa por várias alterações ao longo do tempo. Como aponta Meira (2019), casos que se tornaram globalmente conhecidos como a falência da Enron, seguida por crises financeiras de grandes proporções, foram responsáveis por trazer luz ao papel dos auditores, bem como levantar novas questões sobre sua eficiência e atribuição de responsabilidades.

A auditoria apresenta, entre vários objetivos, a função primária de aumentar o nível de confiança nas demonstrações financeiras apresentadas aos grupos de interesse de uma organização auditada. Para realizar tal objetivo, o auditor utiliza diferentes ferramentas para garantir a fidelidade dos dados e evitar distorções nas informações, afastando os erros e fraudes. Todas atividades relacionadas ao trabalho do auditor tomam novas e diferentes proporções com o aumento da quantidade de dados a serem incorporados nas análises, o que justifica o campo de estudos da

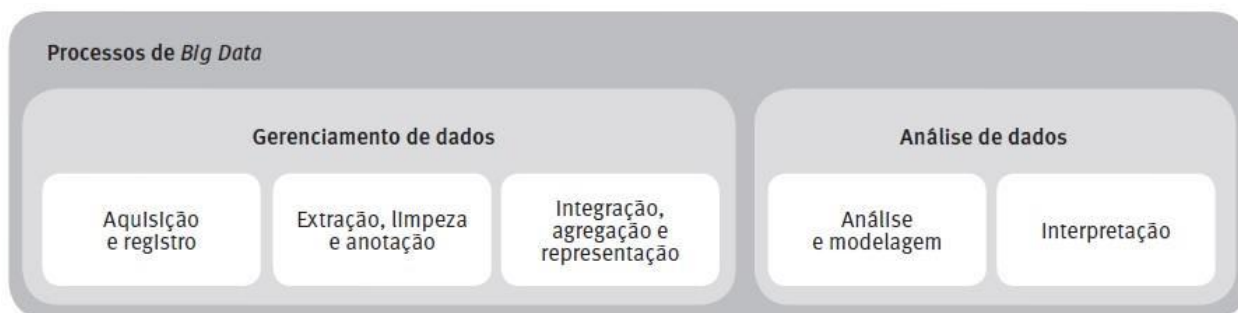
auditoria realizada especificamente com informações provenientes de *Big Data*. (MEIRA, 2019; COSTA; DUTRA, 2014).

2.2.1 Auditoria Digital em *Big Data*

Em meio às novas ferramentas que se apresentam para o manejo e emprego de *Big Data* nas organizações, certamente se destacam aquelas que lidam diretamente com o capital financeiro e com as demonstrações contábeis. A eficácia dos controles operacionais e a verificação de erros e fraudes são setores críticos de qualquer organização, e são tradicionalmente operados pela Auditoria, seja esta interna ou externa. O objetivo das práticas de Auditoria é verificar a existência de erros e, mais importante, desenvolver estratégias que permitam antecipar as necessidades e evitar as não conformidades, prevenindo perdas. A sinalização das falhas, sua análise e encaminhamento de soluções são diretamente derivadas das informações disponíveis ao auditor. Entretanto, com o grande número de dados e transações em meio digital, o abuso ou desvio podem ocorrer sob formas mais difíceis para detecção, tornando-se obscuros em meio ao volume de informações. (TAVARES, 2019).

Ao utilizar *Big Data* como fonte de informações os processos tradicionais para extrair resultados e respostas são modificados. Independentemente das ferramentas pelas quais a organizações opta, pode-se observar alguns processos comuns:

FIGURA 2 – PROCESSOS DE *BIG DATA*



FONTE: CABRERA-SANCHEZ; VILLAREJO-RAMOS (2019, p. 416).

Assim, existem dois processos principais, baseados em gerenciamento e análise dos dados disponíveis. O processo de uso do *Big Data* necessariamente vai desde a aquisição e extração dos dados brutos, passando pela depuração e agregação das informações que serão úteis na análise que se pretende realizar, até a modelagem e interpretação dos resultados encontrados. Para a realização destes processos, existem diferentes técnicas e softwares, que se responsabilizarão também por partes diferentes dos processos, dependendo dos objetivos da análise em *Big Data* que se deseja.

De acordo com Costa e Dutra (2014) o volume crescente de informações deriva do aumento do ritmo de geração e dos formatos cada vez mais diversos sob os quais os dados podem emergir. A quantidade de informações torna-se de difícil manipulação exigindo automação de certos processos, que não podem mais ser realizados apenas pela capacidade humana de processamento. Os autores também colocam que:

Nesse cenário, os mecanismos tradicionais de comunicação de resultados, como as demonstrações anuais, competem com informações financeiras que são divulgadas na Internet e em outros meios, muitas vezes de forma quase simultânea à ocorrência das respectivas transações. Isto favorece decisões mais céleres e contribui para dinamizar o funcionamento da economia e dos governos. (COSTA; DUTRA, 2014, p. 55).

Em tal ambiente, as informações precisam ser cada vez mais confiáveis e, além de demonstrarem com acurácia o passado, precisam apresentar também soluções que permitam prever eventos futuros, facultando melhores decisões na situação atual. De uma forma geral, pode-se considerar que "informações não são úteis se não forem confiáveis" (COSTA; DUTRA, 2014, p. 55).

Costa e Dutra (2014, p. 58) apresentam então o conceito de *Big Data Analytics*, que pode ser definido como:

[...] a aplicação de métodos estatísticos e outras técnicas analíticas sobre dados de transações, informações financeiras e diferentes fontes de dados internas ou externas à organização, com o propósito de extrair conhecimentos a partir do histórico de eventos passados, acompanhar e reagir aos acontecimentos presentes de forma tempestiva ou mesmo prever possíveis desdobramentos futuros a partir dos dados disponíveis.

Com a aplicação de tecnologia ao papel do auditor, ele pode realizar análises mais rápidas e eficientes, referentes a um maior número de dados, e é importante ressaltar a limitação na capacidade cognitiva do auditor humano na transição da auditoria tradicional para a auditoria em *Big Data*. Como a pesquisa de Ahmad (2019) mostra, o auditor humano apresenta limitações nos processos de memória, além de uma série de restrições cognitivas na análise de dados, principalmente dificuldades para operar com sobrecarga de informações, reconhecimento de padrões em grandes grupos de dados, identificação da informação relevantes em meio a dados menos importantes e dificuldades para lidar com a ambiguidade de sentidos na definição das informações. Estas limitações não podem ser compensadas pelo emprego de um número maior de auditores (pois todos teriam as mesmas limitações humanas, mais tempo de investigação ou maior esforço cognitivo, e somente o emprego de ferramentas computacionais permite a atuação do auditor com *Big Data Analytics*.

O alinhamento da auditoria com a velocidade de surgimento dos dados também aponta para um novo caráter neste serviço, que se aproxima cada vez mais de uma prestação de contas em tempo real, com a transparência organizacional sendo auferida continuamente, e não apenas dentro de apresentações periódicas, como os relatórios financeiros anuais. O *Big Data Analytics*, além de possibilitar a prestação constante de informações essenciais, também identifica e analisa padrões, dando visibilidade a anomalias dentro de um grande volume de dados, e extraíndo informações que podem levar à correção precoce de trajetórias indesejadas, além de diminuir riscos. (COSTA; DUTRA, 2014; TAVARES, 2019).

2.2.2 Riscos e Possíveis Benefícios

Os estudos sobre a implementação da automatização digital para *Big Data* nos processos de auditoria têm caráter ainda relativamente recente, e autores como Meira (2019) indicam que em muitos casos os recursos digitais disponíveis não são plenamente empregados para tarefas complexas ou utilizados de forma abrangente, o que limita a investigação aprofundada sobre vantagens e benefícios desta modalidade

de auditoria. A autora, entretanto, extraiu dos próprios auditores as impressões a respeito da implementação da auditoria digital.

De acordo com a pesquisa da autora, a competência das novas gerações para lidar com tecnologia é um fator facilitador na utilização de processos automatizados digitais, tendo a vantagem de diminuir o tempo de implementação e, adicionalmente, contar com um corpo de profissionais que podem também reprogramar e adaptar o sistema, se necessário. O advento gradual, porém inevitável, da auditoria digital tende a formar uma geração de novos auditores que já contarão com tais habilidades e uma maior integração com os processos automatizados. Porém, nesta fase de transição, este também se conforma como um risco para os profissionais mais rígidos, menos confortáveis com o uso da nova tecnologia. É possível que as demandas ligadas ao perfil do auditor sejam modificadas de forma crítica, retirando do mercado aqueles profissionais que não conseguem se adaptar. Para a autora:

Em suma, a tendência evolucionar da profissão é óbvia. Aqueles procedimentos que há alguns anos eram utilizados, hoje são obsoletos, e a auditoria, não só deve, como tem de caminhar no sentido das novas tecnologias para acompanhar os desenvolvimentos dos seus clientes e até ir mais para além disso. Aos auditores resta-lhes alargar horizontes, adquirir novas competências e contribuir com o sentido crítico e o julgamento que tão bem os caracterizam. (MEIRA, 2019, p. 45).

Os auditores também identificam como grande benefício a capacidade dos sistemas de análise de *Big Data* para lidar com um volume praticamente ilimitado de dados, sendo que a auditoria não precisa mais ocorrer por amostras, e pode analisar continuamente todos os dados disponíveis, passando de um modelo retroativo de análise para um modelo preditivo. Na prática isso significa que os eventuais problemas ou desvios podem ser detectados em tempo real, possibilitando a correção precoce de falhas, em comparação com a auditoria tradicional. Porém, o risco que esta vantagem traz está relacionado às complexidades das novas análises, o que pode afastar o auditor da compreensão total dos processos, levando-o a confiar excessivamente na máquina para realizar suas tarefas de acreditação. (MEIRA, 2019).

Outros benefícios relevantes se situam para os clientes. O emprego de processos automatizados permite que a auditoria como um todo se torne mais célere,

exigindo menos participantes e analistas em sua execução. Para o cliente isso se traduz em redução dos honorários a serem pagos, detecção precoce de problemas e agregação de valor ao serviço contratado. Um risco associado, neste caso, seria o fato de que os processos mais ágeis podem levar os auditores a realizar muitas auditorias simultâneas ou em série, para compensar com aumento de volume a redução de honorários por cliente. Certamente isto não é um problema em si mesmo, mas é possível que em alguns casos a qualidade do serviço prestado (no que se refere aos operadores humanos do processo) sofra com o aumento da carga de clientes. Este fator gera um novo risco e pressão sobre os órgãos reguladores, que confrontam-se com a tarefa de estabelecer novos padrões e regras para os serviços de auditoria em um cenário que ainda não é completamente compreendido. (MEIRA, 2019).

Em um caráter mais restrito, no que concerne à prestação de dados e auditoria fiscal, existe um maior volume de pesquisas que identificam as vantagens e desvantagens dos sistemas digitais. Eckert et al. (2011) abordam o Sistema Público de Escrituração Digital (SPED), utilizado por órgãos como a Receita Federal do Brasil, como um mecanismo que efetivamente configura uma norma reguladora em nível nacional, voltada para a auditoria digital contínua das empresas e organizações. O SPED se insere no campo da Escrituração Fiscal Digital, automatizando tarefas de escrituração de documentos fiscais e outras informações que interessam aos fiscos estaduais e federal, agrupando declarações e livros fiscais e integrando a atuação dos próprios órgãos fiscalizadores.

De acordo com os autores, as principais vantagens deste sistema se referem a:

- 1) redução de custos com tributações e diminuição do tempo de analistas voltados apenas para estas tarefas;
- 2) eliminação na duplicidade dos dados;
- 3) Redução dos custos com armazenagem e impressão de documentos;
- 4) facilidade de consulta de dados;
- e 5) incremento do controle nas operações fiscais e financeiras, devido ao fato das informações serem disponibilizadas em tempo real ao Fisco.

Já os riscos ou desvantagens são: 1) custo de implantação do sistema e com a qualificação de trabalhadores para sua utilização; 2) o acesso do Fisco à informação em tempo real pode gerar transtornos imediatos no caso de, por exemplo, emissão de uma nota fiscal com dados errados, levando a autuações que poderiam ser evitadas caso

houvesse tempo para a correção; 3) insegurança a respeito do manejo dos dados detalhados pelo órgão fiscal. (ECKERT et al., 2011).

O SPED não se configura como uma ferramenta abrangente de auditoria digital ao nível organizacional – pois não atua no controle integral do sistema financeiro e produtivo da empresa – mas ilustra adequadamente os benefícios e riscos da implantação real e em larga escala de um sistema de controle e análise de dados (a gestão eletrônica governamental ou *e-government*), especificamente no cenário brasileiro. De acordo com a pesquisa de Ceolato e Behr (2019), o *e-government* ainda representa uma parcela reduzida das tecnologias aplicadas às atividades contábeis (apenas 2%). Em nível internacional, existem outros instrumentos e processos que são utilizados na auditoria em *Big Data*, explorados no tópico a seguir, e dando destaque posterior ao uso da Inteligência Artificial como recurso em fase de implementação e expansão nas auditorias.

2.2.3 Principais Instrumentos e Processos

O mapeamento dos principais instrumentos e processos utilizados na auditoria digital envolve o agregamento de diferentes tecnologias, que experimentam desenvolvimento e mudanças constantes. A pesquisa de Ceolato e Behr (2019) realizou bibliometria abordando as principais tendências tecnológicas citadas em 57 trabalhos acadêmicos, publicados entre 2010 e 2018, e que se referiam à aplicação de novas tecnologias às atividades contábeis, como demonstra a Tabela 1:

TABELA 1 – TECNOLOGIAS APLICADAS ÀS ATIVIDADES CONTÁBEIS

Tecnologia	2010	2011	2012	2013	2014	2015	2016	2017	2018	Total	%
IA						1	1	3	4	9	16%
XBRL		1		1		1	2		2	7	12%
ERP	1		1	1			1	1		5	9%
Mineração de texto (PLN)	1	1					1	1	1	5	9%
Adoção Tecnologias emergentes	1	1				1			1	4	7%
Machine Learning	1		1				1			3	5%
Resource-Event-Agent (REA)			2		1					3	5%
SIC				1				2		3	5%
Automatização	1								1	2	4%
Big data							1	1		2	4%
Blockchain								2		2	4%
Data Analytics							1		1	2	4%
Mecanismos de busca		1					1			2	4%
Computação em nuvem						1				1	2%
Data mining			1							1	2%
Drones								1		1	2%
e-government						1				1	2%
Social Analytics				1						1	2%
Tagscrowdsourced						1				1	2%
Teacher-lessclass					1					1	2%
Teoria da dominância tecnológica					1					1	2%
Total de Artigos	5	4	5	4	3	6	9	11	10	57	100%

FONTE: CEOLATO; BEHR (2019, p. 9).

Como a tabela demonstra, a Inteligência Artificial (IA) ocupa o lugar principal nas pesquisas, seguida pela *Extensible Business Reporting Language* (XBRL). Ainda segundo os autores, mais de um terço dos trabalhos revisados (35%) se referia a Auditoria como a principal área de aplicação das tecnologias.

Estes dados são corroborados por pesquisadores como Meira (2019), que propõe as técnicas de análise de dados através de IA como particularmente apropriadas para a auditoria, já que podem melhorar a qualidade do trabalho final e acelerar os processos relacionados às escolhas de gestão e também aos de tomada de decisão, sendo tais processos menos propensos a erros. De acordo com a autora, tal fato se deve à demanda de automatização daquelas tarefas que se demonstram de caráter repetitivo que fazem parte do processo de auditoria. Também é importante o

uso da IA porque existe um desafio na agregação e compreensão de grandes volumes de dados estruturados e não estruturados nas análises contábeis e nas auditorias. A pesquisadora também sumariza as principais tecnologias que se referem ao uso da IA, apresentadas no Quadro 2:

QUADRO 2 – TECNOLOGIAS REFERENTES AO USO DE INTELIGÊNCIA ARTIFICIAL

Tecnologia	Definição
<i>Machine learning</i>	Capacidade de um computador “aprender” a partir de dados, supervisionados ou não.
Robótica inteligente	Combinação de <i>robots</i> e IA para desempenhar tarefas mais complexas em comparação com os <i>robots</i> tradicionais.
Processamento de linguagem natural	Compreensão, interpretação e criação de linguagem humana natural escrita.
Redes neurais e <i>deep learning</i>	Máquinas que permitem aos modelos de IA aprender de maneira semelhante à dos humanos, imitando o cérebro.
Análise textual	Análise computacional de textos de forma a torná-lo perceptível para sistemas de computadores.
Agentes virtuais	Pessoas virtuais criadas por computadores que interagem com humanos em diferentes contextos (B2C, C2B e B2B ⁷).
Reconhecimento de voz	Permite decifrar linguagem falada e tratá-la como comandos para um computador ou transformá-la em texto escrito.
Visão computacional	Concede a possibilidade de um computador “ver” imagens como os humanos veem.
Biometria	Análise de características humanas, emocionais e físicas. Utilizado para controlo de acessos e identificação.

FONTE: MEIRA (2019, p. 14).

A autora indica que, dentre as tecnologias citadas, o *Machine Learning* tem sido empregado de forma mais frequente pelas grandes organizações, devido à sua capacidade para lidar com grandes quantidades de dados e detectar padrões em volumes de informação de difícil e lento processamento para o auditor humano. Quanto às outras tecnologias, passam a ocupar espaço crescente, de acordo com o desenvolvimento da atuação específica e condições de segurança de dados que se tornam cada vez mais importantes nas auditorias.

Como demonstrado, a Inteligência Artificial tem sido a tecnologia preferencial nas mudanças que são operadas nas auditorias contábeis. A aplicação da IA neste meio traz também mudanças significativas na atuação do auditor ou consultor, com novos processos que necessitam domínio e aplicação pelos profissionais. Estas

mudanças são referentes a todos os passos da auditoria, desde a fase inicial, que inclui o contato com o cliente, passando pela coleta e análise dos dados e finalizando com as avaliações e relatórios consequentes. A comparação entre a auditoria tradicional e auditoria com IA será exposta em mais detalhes nos três próximos quadros, sendo o Quadro 3 referente à fase inicial.

QUADRO 3 – COMPARAÇÃO ENTRE O PROCESSO DE AUDITORIA TRADICIONAL E AUDITORIA COM INTELIGÊNCIA ARTIFICIAL (FASE INICIAL)

Fase de Auditoria	Processo de Auditoria Tradicional	Processo de Auditoria com IA
Pré-planejamento Fase inicial de conhecimento do cliente e do setor no qual se insere	Auditor analisa o setor do cliente Auditor analisa a estrutura organizacional da empresa, os métodos operacionais e os sistemas contábilísticos e financeiros Auditor estima o nível de risco inicial associado ao cliente	IA recolhe e analisa <i>Big Data</i> provenientes de várias fontes Dados relacionados com a estrutura da organização, os métodos operacionais e os sistemas contábilísticos e financeiros alimentam o sistema de IA IA estima o nível de risco inicial associado ao cliente
Contratação Fase de elaboração do contrato e definição dos honorários	Auditor prepara a carta de compromisso com base no risco estimado do cliente. Cliente e Auditor assinam o contrato	IA utiliza a estimativa do nível de risco inicial para calcular os honorários de auditoria e o número de horas que esta exigiria IA analisa uma base de dados de contratos (previamente analisados) e prepara o contrato para o cliente automaticamente Cliente e Auditor assinam o contrato

FONTE: Adaptado de MEIRA (2019, p. 17).

A fase inicial da auditoria é bastante marcada pela interação presencial com o cliente, para estabelecer confiança mútua e garantir o profissionalismo, sigilo entre as partes, acesso privilegiado aos dados da empresa, entre outros diversos fatores que são “negociados” diretamente no contato entre pessoas, representantes da organização e da auditoria. É interessante observar que a integração da IA à fase inicial não interfere negativamente nas relações humanas entre auditor e cliente, como se poderia supor.

De fato, pode-se considerar que ela até ajuda no relacionamento, posto que libera o auditor para trabalhar no estabelecimento do vínculo amistoso e sadio, enquanto a IA trata da parte automatizada da análise dos dados.

Assim, por exemplo, quando o auditor apresenta o nível de risco associado ao cliente como argumento para os honorários propostos, este nível de risco está desvinculado da avaliação pessoal do próprio auditor, que poderia sofrer algum tipo de pressão pelo cliente para que reavaliasse o risco, pressupondo fatores subjetivos no que foi apresentado.

Quando o nível de risco é analisado pela IA não é possível manejar a subjetividade na métrica, e o auditor pode argumentar que a IA se baseia em fatores totalmente objetivos para estimar o risco e as horas de trabalho. Por este motivo garante-se ao cliente que está sendo apresentando um orçamento sem viés e sem superestimativa de horas, evitando a suspeita de que o auditor projeta ganhos financeiros injustificados no processo.

QUADRO 4 – COMPARAÇÃO ENTRE O PROCESSO DE AUDITORIA TRADICIONAL E AUDITORIA COM INTELIGÊNCIA ARTIFICIAL (FASE EXECUÇÃO)

Fase de Auditoria	Processo de Auditoria Tradicional	Processo de Auditoria com IA
<p>Avaliação dos controles internos e identificação dos fatores de risco</p> <p>Fase importante para o planejamento da auditoria</p>	<p>Documentar na pasta de trabalho a informação obtida para entendimento dos controles (fluxogramas, questionários, narrativas, visitas às instalações)</p> <p>Auditor reúne toda esta informação e utiliza o seu julgamento profissional para identificar fatores de risco</p> <p>Entendimento do sistema de controle interno para determinar a extensão, natureza e o <i>timing</i> dos testes substantivos</p>	<p>Informação fornecida pelo cliente (fluxogramas, narrativas e questionários preenchidos) é introduzida no sistema de IA, que usa o reconhecimento de imagem e mineração de textos para análise</p> <p>Em vez de visitas físicas, podem ser utilizados drones para capturar imagens de vídeo que serão posteriormente analisadas pelo sistema</p> <p>IA utiliza o reconhecimento de padrões e métodos de visualização para identificar fatores de risco</p> <p>IA agrega toda a informação de forma a identificar risco de fraude e atos ilegais</p>
Avaliação do risco de controle	Análise das políticas e procedimentos de controle interno da empresa	Sistema de monitorização contínua (baseado em IA) examina a totalidade

Fase de análise do sistema de controle interno do cliente e da sua implementação	Avaliação do risco para cada componente Teste dos controles Reavaliação do risco Documentação dos testes de controles	dos registros e reporta quaisquer violações dos controles IA utiliza mineração de processos para verificar se o sistema de controle interno, mais do que bem desenhado, está implementado corretamente Logs gerados automaticamente para garantir a integridade dos dados e prevenir a sua falsificação
Testes substantivos Fase na qual a diferença entre os dois processos (tradicional vs. IA) é mais pronunciada	Testes periódicos por amostra Testes de detalhe de uma amostra de transações Testes de detalhe de entradas (em determinado momento) Procedimentos analíticos	Qualidade dos dados e da evidência obtida é validada sempre que recolhida (eventualmente em tempo real) IA examina a proveniência dos dados Testes de detalhe para 100% das transações numa base contínua Testes de detalhe para entradas contínuas (em todos os momentos) Reconhecimento de padrões, visualização, <i>benchmarks</i> e métodos de detecção de <i>outliers</i> aplicados continuamente, aliados aos procedimentos analíticos

FONTE: Adaptado de MEIRA (2019, p. 17-18).

A fase de execução torna ainda mais explícita a diferença introduzida pela IA no processo de auditoria. A avaliação dos riscos ocorre com grande nível de automatização, e não depende mais do *feeling* ou *know-how* do auditor para estabelecimento objetivo de uma visão crítica da organização. Isso não significa que o auditor está dispensado de seu papel, mas sim que ele empregará seu tempo de forma mais qualitativa, utilizando sua experiência para legitimar e organizar os fatores que foram encontrados pela IA. A vantagem da IA nesta fase do processo é que ela pode coletar e processar dados extensivos, indexando pontos de risco que talvez tivessem “escapado” sob uma análise humana inicial.

A avaliação e análise do sistema de controles por IA também traz vantagens expressivas de tempo e credibilidade ao trabalho do auditor. Na análise dos controles internos a IA pode vasculhar todos os processos, em busca de indicadores de falhas

mínimas, prevenindo problemas futuros ou apontando melhorias a serem feitas internamente. Além disso, a geração automática de *logs* ao fim das análises garante ao cliente que não houve manipulação indevida dos resultados, evitando a suspeita de que determinados setores ou departamentos estão sendo subavaliados ou recebendo tratamento diferente na auditoria.

A fase de testes, que costuma encerrar a execução e caracteriza o maior volume de dados a serem analisados, mostra com clareza a vantagem do uso da IA na auditoria. A limitação cognitiva humana sempre determinou que os testes fossem realizados por amostras, com entradas (preferivelmente) aleatórias para detectar falhas de maneira fortuita. O aumento do número de amostras e da periodicidade de recolhimento das entradas determina a qualidade da auditoria, mas este é um fator que é limitado pela equipe de auditores envolvidos, pelo esforço dedicado à coleta, qualificação e sistematização de dados, e pelo tempo dedicado às análises.

Com a utilização da IA os limites são drasticamente reduzidos em comparação à capacidade humana, e 100% das entradas podem ser analisadas, garantindo cobertura total dos dados disponíveis. A IA também se responsabiliza pela qualificação das entradas, examinando a proveniência dos dados e assegurando origem. Finalmente, a capacidade da IA para reconhecer padrões em diferentes escalas gera rapidamente visões analíticas sobre a organização que seriam extremamente custosas (em termos financeiros e de tempo) para produção por uma equipe de auditores humanos.

QUADRO 5 – COMPARAÇÃO ENTRE O PROCESSO DE AUDITORIA TRADICIONAL E AUDITORIA COM INTELIGÊNCIA ARTIFICIAL (FASE FINAL)

Fase de Auditoria	Processo de Auditoria Tradicional	Processo de Auditoria com IA
<p>Avaliação da evidência obtida</p> <p>Fase para verificar a qualidade e credibilidade da prova obtida</p>	<p>Auditor deve avaliar a suficiência, clareza e aceitação da evidência recolhida. Consequentemente, o auditor pode recolher mais evidência ou desistir do compromisso.</p>	<p>Passa a fazer parte da fase anterior dada a importância de garantir a qualidade dos dados antes da execução dos testes substantivos</p>
<p>Relatório de</p>	<p>Auditor reúne a informação obtida nas</p>	<p>IA utiliza um modelo preditivo para</p>

auditoria	fases anteriores para emitir o relatório	estimar os vários riscos identificados
Última fase do processo de auditoria que consiste na emissão da opinião do auditor	Relatório é taxativo: limpo, qualificado, adverso, etc.	Relatório pode ser contínuo (por exemplo numa escala de 1-100) em vez de taxativo

FONTE: Adaptado de MEIRA (2019, p. 18).

Com o emprego da IA a fase final da auditoria, além de ser implementada de forma mais ágil, também pode deixar um legado mais substantivo ao cliente. Como apresentado nos Quadros 4 e 5, a IA qualifica e examina a procedência de todos os dados quando estes são coletados e o sistema é alimentado, apresentando um conjunto de dados extremamente consistente e confiável para receber o aval final do auditor. A certificação ao término torna-se mais rápida e padronizada, sendo que o processo sofre acreditação desde o início, já que dados irregulares ou inadequados gerariam advertências e visualizações de falhas sistêmicas durante o processo, não sendo mais necessário o reexame pós-teste.

Logicamente a figura de responsável técnico não deixa de existir, mas o auditor principal pode agora assinar o documento final com mais segurança a respeito dos resultados. Os documentos finais, principalmente quando estes se referem às grandes organizações, tradicionalmente demandam que o auditor principal certifique o trabalho realizado por grandes equipes de auditores e assistentes (e, portanto, mais sujeito a erros). A fragilidade representada pelo emprego de muitas pessoas na análise de um mesmo conjunto de dados deixa de existir com o uso da IA.

E, finalmente, o relatório final gerado pela IA pode ser apresentado em escala progressiva, resumizando ou pontuando todas as métricas que geram um score final. Com esta visão mais detalhada, baseada em scores e não em julgamentos taxativos, a interpretação dos dados está aberta para reavaliação por qualquer interessado nos riscos da empresa, e o auditor não precisa constantemente “dar explicações” sobre seus critérios para uma avaliação favorável ou desfavorável.

3 PROCEDIMENTOS METODOLÓGICOS

Esta seção apresenta a metodologia, os materiais e os métodos utilizados para que os objetivos fossem alcançados.

3.1 METODOLOGIA

O estudo proposto nesta monografia é baseado no método de *revisão assistemática ou narrativa de literatura* que, de acordo com Vosgerau e Romanowski (2014), são análises qualitativas abrangentes da produção acadêmica disponível. O objetivo da revisão assistemática é ter uma visão abrangente sobre um tema ou pergunta, observando quais são as produções disponíveis para apontar novos caminhos de estudo ou investigação. Também, como definido por Rother (2007), considera-se que a revisão narrativa é aquela que busca responder a um questão ampla, para a qual existe dificuldade na especificação de fontes, inicialmente. Em contraste com as revisões sistemáticas, a revisão narrativa ou assistemática não apresenta obrigatoriamente uma avaliação criteriosa ou reproduzível das referências, já que ela é baseada fundamentalmente em uma avaliação crítica do autor a respeito daquilo que apresenta coerência para levantar o estado da arte a respeito do tema de pesquisa.

Mesmo dentro da concepção assistemática, esta revisão procura apresentar o detalhamento da sua estrutura de busca inicial, mostrando também como ocorreu a seleção das referências analisadas. Diferente de uma revisão *sistemática* de literatura, não há objetivo de reprodução ou replicação da busca, mas considera-se que é necessária transparência do processo seletivo, destacando a literatura disponível sobre o tema e registrando-a como memória no caso de futura ampliação de estudos dentro deste assunto e com objetivos similares. Note-se ainda que a busca inicial retornou poucos resultados, o que inviabilizaria a composição deste estudo apenas através de uma revisão sistemática. Sendo assim, a opção de revisão assistemática, com avaliação crítica do material encontrado pelo autor, foi utilizada como complemento para a fundamentação teórica.

3.2 MATERIAIS

Considerando os objetivos deste trabalho, foi realizada revisão de literatura para mapear as discussões mais recentes a respeito de auditoria no cenário de *Big Data*. As referências discutidas foram selecionadas por conveniência e afinidade com os objetivos geral e específicos estabelecidos na Introdução. Para o mapeamento inicial das referências disponíveis foi realizada busca no Portal de Periódicos CAPES/MEC (<https://www.periodicos.capes.gov.br>) e no SciELO (<https://scielo.org>). As *strings* utilizadas na busca decorrem de pesquisa livre sobre o tema na *search engine Google* (<https://www.google.com>), que definiram como de maior ocorrência nos trabalhos acadêmicos as expressões “auditoria”, e “*Big Data*”. Também foram observadas ocorrências para “auditoria digital”, mas esta categoria foi considerada específica demais como restritor da pesquisa, o que poderia retornar resultados muito escassos.

A busca de referências foi realizada em 17 de maio de 2020 com a configuração básica ["*auditoria*" AND "*big data*"] e os resultados foram restritos aos 10 últimos anos de produção (2010-2020). Foram selecionadas apenas referências de periódicos revisados por pares, com os idiomas espanhol, inglês e português. Foram eliminadas as referências sem texto completo disponível ou com restrições de acesso no momento da consulta. O Quadro 6 apresenta os resultados da busca inicial.

QUADRO 6 – PROCESSO DE BUSCA DAS REFERÊNCIAS

Base de dados	Resultados totais para a <i>string</i> de busca	Resultados totais para referências de <i>periódicos revisados por pares, entre 2010-2020 e idiomas inglês, português e espanhol</i>	Resultados após remoção de referências não artigo, sem texto completo, ou com restrições de acesso
Periódicos CAPES	31	27	24
SciELO	06	06	06
TOTAL	37	33	30

FONTE: O Autor (2020)

Como o quadro demonstra, a busca inicial revelou 37 referências, dentre as quais foram removidas aquelas com erros de cruzamento de informações resultantes da busca pela base de dados ou com restrições de acesso de algum tipo (como base de dados indisponível ou erro de endereço de acesso), além de aplicação dos critérios iniciais de seleção (idiomas e ano de publicação).

A seleção inicial foi reduzida para 30 resultados aos quais foram aplicados critérios de elegibilidade, através da leitura dos resumos dos textos que se encontravam disponíveis. O principal critério de inclusão utilizado foi a apresentação de relação explícita entre pesquisas em *Big Data* e sua aplicação em qualquer forma de auditoria. Foram também aplicados os seguintes critérios para exclusão de referências que apresentassem:

- Estudos ou pesquisas genéricas apenas com o tema *Big Data*;
- Estudos ou pesquisas genéricas apenas com o tema auditoria;
- Resultados ou sumários de congressos ou eventos versando sobre Gestão da Informação ou Tecnologia, mas sem relação com auditoria;
- Ausência de relação entre temas de auditoria com análise de grandes bases de dados (mesmo sem citar o termo “*Big Data*”);

Os critérios de inclusão e exclusão reduziram o total de referências para 4. A leitura mais aprofundada dos textos integrais resultantes indicou quais seriam as referências a serem empregadas nesta pesquisa. Desta última seleção, baseada na leitura crítica de cada texto integral, restaram 3 referências. O Quadro 7 apresenta o processo de seleção entre os resumos totais encontrados e as referências que foram de fato utilizadas.

QUADRO 7 – PROCESSO DE SELEÇÃO DAS REFERÊNCIAS

Total de referências	Eliminação de duplicatas e erros de cruzamento nas bases de dados	Aplicação dos critérios de elegibilidade	Referências empregadas na revisão crítica
----------------------	---	--	---

30	28	04	03
----	----	----	----

FONTE: O Autor (2020)

O Quadro 8 relaciona as 3 referências que foram empregadas na revisão de literatura realizada, apresentando os autores, data de publicação e o título do trabalho.

QUADRO 8 – LISTA DE REFERÊNCIAS SELECIONADAS E EMPREGADAS

Ref	Autor/Ano	Título
01	AHMAD (2019)	<i>A systematic review of the role of Big Data Analytics in reducing the influence of cognitive errors on the audit judgement</i>
02	ESTEBAN-NAVARRO; GARCÍA-MADURGA (2019)	<i>Estudio exploratorio sobre la percepción de la gestión de la Inteligencia Competitiva por los directivos en empresas aragonesas sin una práctica sistematizada</i>
03	PEDROSA; LAUREANO; COSTA (2015)	<i>Motivações dos auditores para o uso das Tecnologias de Informação na sua profissão: aplicação aos Revisores Oficiais de Contas</i>

FONTE: O Autor (2020)

O número final de 3 referências foi considerado insuficiente para a revisão de literatura realizada nesta monografia. Para complementar as referências da fundamentação teórica foi realizada também busca livre na *search engine Google* e na sua subdivisão *Google Scholar* (<https://scholar.google.com>). Com estas buscas, realizadas entre diferentes períodos de estruturação da monografia, as referências foram progressivamente adicionadas por convergência com o tema da pesquisa, considerando fatores como número de citação por outros autores, indexação em revistas especializadas, restrição ao período de publicação 2010-2020 e apresentação nos idiomas espanhol, inglês ou português.

3.3 MÉTODOS DE ANÁLISE

Todas as referências foram analisadas para busca dos temas definidos nos tópicos da fundamentação teórica, com extração e distribuição dos dados relevantes

das pesquisas ao longo do texto. Os conceitos encontrados nos autores originais foram articulados com o tema da pesquisa e também entre os próprios trabalhos analisados, buscando oferecer múltiplas visões a respeito do processo de auditoria digital em *Big Data*. A fundamentação progride de uma visão geral de Big Data e auditoria digital até a identificação dos principais processos e instrumentos disponíveis atualmente para este tipo de auditoria, compondo um contínuo que leva da visão teórica sobre o tema até sua implementação de fato nos processos atuais da auditoria.

Posteriormente, os conceitos explorados na fundamentação teórica são integrados e analisados de forma mais crítica no capítulo de discussão, que resgata os achados anteriores da pesquisa e agrega novos autores para refletir sobre a aplicação prática contemporânea dos processos de auditoria digital, com o objetivo de criar um quadro ilustrativo de como esse processo, ainda incipiente em muitos casos, vem ocorrendo no cenário internacional.

4 DISCUSSÃO

A revisão da literatura realizada na fundamentação teórica deste trabalho mostrou diferentes riscos e benefícios no emprego das ferramentas digitais de auditoria, e que estas variáveis se tornam ainda mais pronunciadas em relação direta com o volume de dados a serem analisados. Sendo assim, a análise e auditoria de dados massivos – o *Big Data Analytics* – expõe de forma crítica as diferenças de atuação e resultados entre a auditoria digital e a tradicional.

Embora ainda exista certa resistência à adoção de tais recursos por alguns escritórios ou auditores independentes, e mesmo considerando os problemas que a confiança excessiva na máquina pode gerar, pode-se dizer com razoável grau de certeza que a utilização de recursos informatizados é um caminho sem volta, pelo menos para aqueles auditores que desejam atender às demandas de médios e grandes clientes. Esta tendência é ilustrada pela crescente adoção de novas tecnologias de IA para auditoria pelos grandes escritórios como *PricewaterhouseCoopers*, *KPMG*, *Deloitte* e *Ernst&Young*, que deixam clara a necessidade de atender às novas expectativas de informação relevante dos investidores e gestores de grandes corporações. (BRANDAS; MUNTEAN; DIDRAGA, 2018).

Deve-se notar também, analisando a literatura sobre o tema, que as práticas de auditoria digital passam por uma evolução acelerada e seu acompanhamento pode ser bastante difícil. A tecnologia integrada aos processos de auditoria torna-se rapidamente complexa, principalmente quando se trata de dados não estruturados e da combinação de diferentes fontes de informação, empregando simultaneamente recursos como a IA, ERP, mineração de texto, aprendizado de máquina e mineração de dados (entre outros), com disponibilidade ainda em ascensão e crescente implementação nos softwares de apoio à auditorias (CEOLATO; BEHR, 2019).

O aprendizado de máquina (*machine learning*) integrado à IA constitui um dos exemplos mais importantes da obrigatoriedade de mudança no posicionamento e atuação do auditor. Como apontado por Brandas, Muntean e Didraga (2018), atualmente a maior parte das transações financeiras e dos documentos são gerados, processados, armazenados e transmitidos de forma eletrônica, e cada vez menos

documentos possuem uma cópia, registro físico ou impresso. A existência de tais registros e dados apenas em formato digital aumenta muito os riscos da auditoria para os responsáveis técnicos, pois o método de auditoria por amostragem torna-se mais irrelevante e inseguro em proporção direta com o aumento do volume de dados. Some-se a esta observação o fato de que muitos registros e dados importantes podem ser não estruturados, ou seja: são figuras, imagens, fotos de documentos, textos processados em formato digital, dados numéricos relevantes no meio de descrições textuais, etc.

Estas fontes de dados não podem ser ignoradas, mas ao mesmo tempo podem exceder a capacidade de análise de um corpo de auditores em um determinado período de tempo. O uso da IA permite que sejam empregados algoritmos pré-definidos nas análises, que detectam e categorizam todas as fontes e dados disponíveis, apresentando um conjunto de dados manejável em escala humana para análise. O tempo de cada auditor é otimizado, e garante-se que nenhum dado foi ignorado por falta de tempo ou de recursos humanos.

O aprendizado de máquina leva os mesmos processos a um passo adiante; constitui-se em sistema que, assim como a IA, possui algoritmos para seleção de dados, categorização de grupos de informações e automatização de tarefas, mas que ajusta (e até mesmo cria) os próprios algoritmos de acordo com a interação com outros sistemas, incluindo operadores humanos. Assim, além de cumprir as tarefas previamente definidas, o sistema pode gerar e sugerir novos processos, posto que não está limitado à sua programação inicial (PINTO, 2020). Embora ainda seja uma área incipiente, o uso de *machine learning* nos sistemas de suporte à auditoria tem potencial para aumentar enormemente a eficiência e diminuir significativamente os riscos. Se estas expectativas de fato se comprovarem, então a auditoria tradicional possivelmente deixará de existir.

O estudo realizado por Hooda, Rana e Bawa (2018) traz um exemplo da aplicação de algoritmos de *machine learning* na detecção de empresas com risco de fraude financeira. Foram empregados na análise por algoritmos os dados disponíveis em relatórios anuais de 777 empresas, de 14 setores diferentes. Através da aplicação de diferentes modelos de IA os pesquisadores encontraram sistemas que chegam a

93% de acerto na identificação de empresas com suspeita de fraude. A identificação da suspeita de fraude é um importante mecanismo seletivo para os auditores, que normalmente precisam realizar um extenso trabalho prévio de coleta e análise de dados internos, para decidir quais escritórios ou filiais serão fisicamente visitados para verificação dos relatórios e informações financeiras. O desenvolvimento de sistemas eficientes, que realizam rapidamente a seleção dos escritórios suspeitos, diminui significativamente o trabalho dos auditores, empregando os recursos disponíveis de forma mais eficiente em um processo de auditoria.

A incorporação constante destes recursos aos processos de auditoria indica que já existem e aumentarão as demandas de domínio técnico de diferentes softwares pelos profissionais auditores. Conhecimentos especializados e jargão da Informática, que até pouco tempo atrás faziam parte do ferramental apenas de desenvolvedores de TI, passam a constituir parte da formação profissional de estudantes de Contabilidade e profissionais de Controladoria, implicando também em atualização e aprendizado constante pelos profissionais que já estão posicionados no mercado. A insistência no uso dos métodos e técnicas tradicionais – quando ignoram completamente as novas tecnologias – levam a resultados ambíguos nas auditorias, dificuldades maiores na análise dos dados e entrega de resultados pouco relevantes aos clientes.

Estes fatos interferentes na auditoria mudam o caráter do profissional auditor e também das tarefas que ele desempenha, levantando perguntas sobre a legitimidade de colocar os processos automatizados em primeiro lugar e, depois, fazer o profissional “correr atrás” destas mudanças. Como já citado em diferentes pontos deste texto, a eficiência destes sistemas é inegável e imbatível em termos de abrangência de dados e velocidade de análise. O contraste de desempenho entre máquina e humano leva ao privilégio do sistema automatizado para a execução da tarefa, e sucessivas ações passam a ser atribuídas ao sistema que, pela sua crescente complexidade, restringe progressivamente o número de humanos que o compreendem e podem operá-lo.

Porém, como ainda não há aplicação plena de tais sistemas, muitas dúvidas existem sobre quais decisões importantes poderiam estar sob responsabilidade da IA. A revisão da literatura deixou claro que ainda não existem parâmetros éticos para definir especificamente quais decisões podem ser tomadas pelo sistema, e quais necessitam

de aprovação ou reflexão humana. O problema se aprofunda no caso do *machine learning* pois, como Pinto (2020) descreve, dependendo da maturidade ou complexidade do aprendizado realizado pelo sistema os seus parâmetros de decisão não estão abertos ou disponíveis para exame técnico, e auditar ou examinar o seu código-fonte não traria indicações claras a respeito de como o sistema realmente atua e decide, depois de seguidas iterações.

Em aplicações práticas de sistemas de IA aos procesos de auditoria observa-se que os vieses e limitações técnicas dos desenvolvedores (e também dos operadores dos sistemas) podem interferir negativamente na atuação da IA gerando inequidades e fragilidades nos resultados da auditoria. Sendo assim, as aplicações atuais de tais sistemas indicam que existe a necessidade de estabelecer parâmetros restritos de avaliação do desempenho dos sistemas automatizados, bem como mecanismos de regulação que minimizem seus possíveis riscos sociais (RAJI et al., 2020).

Chamou também a atenção nesta revisão de literatura os novos potenciais caminhos de coleta de informação e emergentes decisões que dão uma nova camada de complexidade ao processo da auditoria. A incorporação da análise de *Big Data* indica maneiras diferentes de levantar os riscos da empresa, propondo uma série de novos critérios que não eram rotineiramente empregados nas auditorias tradicionais. Com os processos automatizados o auditor se depara com a possibilidade de tratar grupos de dados que poderiam ser desconsiderados anteriormente. Sun e Vasarhelyi (2018) indicam que o auditor deve olhar não apenas para os dados internos da empresa, mas também para notícias sobre a empresa na mídia, ou para menções à empresa nas redes sociais. O uso da IA e de *machine learning* permite que grandes séries de dados textuais, como *posts*, comentários, artigos e *tweets*, sejam explorados para o levantamento de metadados referentes à organização auditada.

O levantamento de metadados que dizer que a IA tem condições de analisar grandes blocos de informações, de diferentes fontes, para derivar informações como tópicos que se repetem, palavras-chave, menções da empresa em conjunto com outras organizações, eventos positivos ou negativos que marcaram a empresa na mídia, anomalias e até mesmo sentimentos que circulam em relação à empresa. A extração e organização destes metadados pode, por exemplo, informar os auditores a respeito dos

níveis de insatisfação dos clientes com a empresa, auxiliando na estimativa dos riscos de litígio que podem ocorrer. Os metadados podem ser similarmente empregados nas estimativas de riscos do negócio, de controle interno ou de fraude na gestão. (SUN; VASARHELYI, 2018). E este potencial se refere apenas à análise textual, sendo que a IA e *machine learning* podem atuar com diferentes tipos de fontes, inclusive análise de voz e imagens, como já citado anteriormente.

O aumento da abrangência de dados disponíveis para análise coloca novas questões na implementação dos próprios sistemas, com problemas que emergem de tais aplicações. Veale, Kleek e Binns (2018) entrevistaram 27 profissionais que atuam com aprendizado de máquina nos escritórios internacionais da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), e concluíram que existe uma relevante disparidade entre a pesquisa atual dos sistemas de *machine learning* e as realidades institucionais e organizacionais de aplicação destas tecnologias. Entre as observações dos entrevistados encontrou-se que existe excesso de confiança e também falta de confiança nos processos automatizados, o que prejudica sua aplicação efetiva; no caso de excesso de confiança, muitos processos deixam de ser certificados por um analista humano, diminuindo a legitimidade do uso do sistema quando ocorrem erros ou desvios. Por outro lado, quando há falta de confiança no sistema, os processos nunca são inteiramente implementados e as análises permanecem parciais ou de pouca utilidade.

Também existem dificuldades relacionadas com a transparência dos processos automatizados. Principalmente quando se trata de dados públicos sensíveis (como os sistemas fiscais), existe um equilíbrio difícil entre tornar os algoritmos acessíveis ou manter o sistema opaco aos usuários. Isto se justifica porque, ao abrir todos os dados, pode ocorrer manipulação do sistema ou dos dados, para beneficiar aqueles que estão sendo auditados ou burlar mecanismos mais rigorosos de análise. Além disso, os autores indicam que muitos casos de implementação envolvem o acompanhamento por analistas humanos, em todas as fases, porque todos os sistemas automatizados, pelo menos por enquanto, ainda precisam de *input* humano constante para garantir que o aprendizado da máquina siga critérios éticos e regulamentações estritas. Como a máquina não realiza julgamentos, apenas aplica algoritmos às séries de dados, nem

sempre os padrões e informações encontrados atendem às necessidades dos processos de auditoria, e máquina não entende intrinsecamente quais tipos de dados são sensíveis e precisam ser preservados ou ocultados.

Finalmente, os autores indicam que os processos que incluem *machine learning* ainda estão em fase preliminar, e que há necessidade de melhorias de análises e aprimoramentos quando tais sistemas são implementados. Em muitos cenários políticos, governamentais ou sociais estes esforços podem significar muitas horas de envolvimento de especialistas, o que pode dificultar ou impossibilitar os trabalhos necessários ao desbravamento deste setor. Além disso, é uma área de estudo e aplicação que envolve intensa interdisciplinariedade, e colaboração estreita entre os operadores da tecnologia, os usuários ou outros indivíduos afetados por seus resultados (VEALE; KLEEK; BINNS, 2018). Com o aumento da escala de análise, parece que também aumenta o número de pessoas envolvidas na sua aplicação prática. E, na realidade organizacional ou institucional, nem sempre pode haver disponibilidade de recursos financeiros e humanos para atender a estas demandas.

Em síntese, esta discussão conclui que o uso de *Big Data* em auditoria apresenta enorme potencial, trazendo vantagens e mudanças significativas para os profissionais e clientes. Ao mesmo tempo, também gera alterações em vários critérios e passos dos processos tradicionais de auditoria. Entretanto, foi observado que o emprego de tais processos automatizados para análise de grandes séries de dados ainda encontra-se no início, e que existem muitas questões a serem resolvidas, desde o nível técnico até o nível ético, envolvendo principalmente qual é o nível adequado de participação humana no processo de auditoria, e como deve ser a regulação e avaliação dos próprios sistemas automatizados. Fica claro que, na busca de soluções para tais perguntas, é importante manter uma perspectiva humana na auditoria, na qual a máquina sirva aos profissionais e às empresas, mas sem tomar o lugar do auditor.

5 CONSIDERAÇÕES FINAIS

O estudo apresentado neste trabalho delineou alguns pontos principais em relação às auditorias realizadas com *Big Data*, e que atendem aos objetivos descritos inicialmente. A primeira observação é de que o cenário brasileiro enfrenta dificuldades similares às de outros países na adoção das auditorias digitais e informatizadas. Os principais entraves são referentes à implementação de novas tecnologias – que podem ser muito caras para uma empresa pequena ou média – e à formação de recursos humanos, que são compostos por uma massa jovem (que não aprende sobre o tema na sua formação inicial) e pelo grupo de consultores veteranos (que apresenta resistência à formação complementar).

Deve-se considerar, entretanto, que as corporações brasileiras são estimuladas a adotar as novas tendências de auditoria por algumas iniciativas governamentais, como a implementação do SPED em nível nacional. A obrigatoriedade da prestação de contas contínua à Receita Federal leva muitas empresas a se adequarem à nova realidade das auditorias. Sendo assim, no Brasil, o *e-government* pode ser um importante instrumento de estímulo para que mais empresas e auditorias embarquem no novo modelo baseado em análise de *Big Data* e uso de IA.

Outra observação relevante é que a tecnologia ainda está em franco desenvolvimento e ocorre muita experimentação. Os quadros técnicos de TI e os auditores engajam muitas horas de trabalho na construção e testagem de ferramentas que nem sempre trazem resultados satisfatórios ou superiores à análise humana. Mesmo assim, a literatura revisada sugere que esta aceleração na experimentação de novas tecnologias deve continuar pelo futuro próximo, considerando que existe substancial envolvimento e financiamento tanto por parte dos desenvolvedores (como *IBM* e *Google*) quanto dos grandes escritórios de auditoria. Embora nem sempre sejam constatados resultados qualitativos, o desenvolvimento é bastante estimulado pela capacidade de análise quantitativa, ou seja, a possibilidade que a IA traz de categorizar e analisar dados massivos e não estruturados.

Finalmente observa-se que os processos automatizados, digitais e informatizados discutidos não dispensam a presença e legitimação de um auditor

humano em praticamente todas as fases do processo de auditoria. Seja na definição dos *inputs*, qualificação dos algoritmos de categorização e análise, certificação dos relatórios ou seleção e apresentação final dos resultados, verifica-se que, por enquanto, o auditor humano é figura essencial e não pode ser substituído pelo software. E, ainda que a parte técnica seja eventualmente incorporada pela máquina, ainda assim restará ao auditor humano o papel fundamental de direcionamento ético do processo, posto que a ferramenta ou software não possui condições para classificar os dados e resultados de acordo com os efeitos sociais e implicações morais que este podem ocasionar.

Embora esta pesquisa tenha cumprido seus objetivos, notou-se ao longo do processo que a produção nacional sobre o tema ainda é bastante tímida, com poucos estudos de casos e descrições das aplicações práticas de análises de *Big Data* ou do uso de IA nas auditorias. Mesmo em nível internacional, os resultados em língua inglesa também são bastante restritos e a maioria parece estar contida em *folders* ou encartes eletrônicos de escritórios de auditoria, funcionando como propaganda institucional, mas não como estudos organizados de forma acadêmica. Portanto, esta pesquisa foi restrita no cumprimento de seus objetivos pela escassa bibliografia indexada que foi encontrada, como descrito na parte de metodologia.

A sugestão para pesquisas futuras neste mesmo tema é de que sejam repetidas dentro de intervalos regulares e curtos, posto que as tecnologias que empregam IA e análise de *Big Data* estão em ascensão e mudança constante. Também sugere-se que o mapeamento da pesquisa seja iniciado pelos estudos de caso e aplicações práticas da tecnologia, sendo que a fundamentação teórica pode seguir o que se observa na prática, legitimando ou falseando as hipóteses iniciais a partir do que realmente ocorre nas organizações e processos de auditoria com uso de *Big Data*.

REFERÊNCIAS

AHMAD, F. A systematic review of the role of Big Data Analytics in reducing the influence of cognitive errors on the audit judgement. **Revista de Contabilidade – Spanish Accounting Review**, v. 22, n. 2, p. 187-202, 2019. Disponível em: <<https://dialnet.unirioja.es/servlet/articulo?codigo=6981559>>. Acesso em: 17 mai. 2020.

AJAH, I. A.; NWEKE, H. F. Big Data and Business Analytics: Trends, Platforms, Success Factors and Applications. **Big Data and Cognitive Computing**, v. 3, n. 32, 2019. Disponível em: <[doi:10.3390/bdcc3020032](https://doi.org/10.3390/bdcc3020032)>. Acesso em: 06 mar. 2020.

ALVES, P. M. de A.; FREITAS, A. de O. Ferramentas informatizadas utilizadas na auditoria. **Revista Brasileira de Contabilidade**, 225, p. 79-87, 2017. Disponível em: <<https://pdfs.semanticscholar.org/990a/69131f99d85d472c044143a4674d69893286.pdf>>. Acesso em: 19 jun. 2020.

ANDREU-PEREZ, J.; POON, C. C. Y.; MERRIFIELD, R. D.; WONG, S. T. C.; YANG, G. Z. Big Data for Health. **IEEE Journal of Biomedical and Health Informatics**, v. 19, n. 4, p. 1193-1208, 2015. Disponível em: <<https://ieeexplore.ieee.org/document/7154395>>. Acesso em: 06 jun. 2020.

BRYNJOLFSSON, E.; HITT, L. M.; KIM, H. H. Strength in Numbers: How Does Data-Driven Decision Making Affect Firm Performance? (Documento online). **SSRN**, 2011. Disponível em: <<http://ssrn.com/abstract=1819486>>. Acesso em: 07 mar. 2020.

CABRERA-SANCHEZ, J-P.; VILLAREJO-RAMOS, A. F. Fatores que Afetam a Adoção de Análises de Big Data em Empresas. **RAE**, v. 59, n. 6, p. 415-429, 2019. Disponível em: <[DOI: http://dx.doi.org/10.1590/S0034-759020190607](http://dx.doi.org/10.1590/S0034-759020190607)>. Acesso em: 03 mai. 2020.

CASALINHO, G. D. O. O Impacto do Uso do Big Data na Inteligência Competitiva e na Percepção do Produto pelo Cliente: Desenvolvimento de Proposições de Pesquisa. **ESTUDO & DEBATE**, v. 22, n. 2, p. 154-170, 2015. Disponível em: <<http://www.univates.br/revistas/index.php/estudoedebate/article/download/660/650>>. Acesso em: 07 mar. 2020.

CEOLATO, R.; BEHR, A. **Análise Bibliométrica de Artigos da Área de Sistemas de Informação Contábil e suas Contribuições Relacionadas à Aplicação de Tecnologias Emergentes na Contabilidade**. 20 f. Trabalho de Conclusão de Curso (Bacharelado em Ciência Contábeis) – Universidade Federal do Rio Grande do Sul, 2019. Disponível em: <<https://lume.ufrgs.br/bitstream/handle/10183/198557/001098050.pdf?sequence=1&isAllowed=y>>. Acesso em 10 mai. 2020.

CLAUDIU, B.; DIDRAGA, O.; MUNTEAN, M. Intelligent Decision Support in Auditing: Big Data and Machine Learning Approach. **Proceedings of the IE 2018 International Conference**, p. 425-430, 2018. Disponível em:

<https://www.researchgate.net/profile/Brandas_Claudiu/publication/328858614_INTELLIGENT_DECISION_SUPPORT_IN_AUDITING_BIG_DATA_AND_MACHINE_LEARNING_APPROACH/links/5be701a7a6fdcc3a8dcc125e/INTELLIGENT-DECISION-SUPPORT-IN-AUDITING-BIG-DATA-AND-MACHINE-LEARNING-APPROACH.pdf>. Acesso em: 07 jun, 2020.

COSTA, G. P. C. da; DUTRA, T. A. de G. L. Auditoria financeira na era do *Big Data*: novas possibilidades para avaliação e resposta a riscos em demonstrações financeiras do Governo Federal. **Revista TCU**, n. 131, p. 54-61, 2014 Disponível em: <<https://revista.tcu.gov.br/ojs/index.php/RTCU/article/view/62>>. Acesso em: 26 abr. 2020.

De MAURO, A.; GRECO, M.; GRIMALDI, M. A Formal Definition of Big Data Based on its Essential Features. **Library Review**, v. 65, n. 3, p.122-135, 2016. Disponível em: <DOI: 10.1108/LR-06-2015-006>. Acesso em: 02 mai. 2020.

ECKERT, A.; SANTOS, E. C. dos; MECCA, M. S.; BIASIO, R. Vantagens e Desvantagens da Implantação do Sistema Público de Escrituração Digital em Uma Grande Empresa Industrial. **Revista de Contabilidade e Controladoria**, v. 3, n. 3, p. 82-93, 2011. Disponível em: <http://www.biasio.pro.br/wp-content/uploads/Outros/artigo_cientifica/periodicos/vantagens_e_desvantagens_da_implantacao_rev._rcc_2011.pdf>. Acesso em: 30 mai. 2020.

ESTEBAN-NAVARRO, M. A.; GARCÍA-MADURGA, M. A. Estudio exploratorio sobre la percepción de la gestión de la Inteligencia Competitiva por los directivos en empresas aragonesas sin una práctica sistematizada. **Revista Española de Documentación Científica**, v. 42, n. 3, 2019. Disponível em: <<https://doi.org/10.3989/redc.2019.3.1607>>. Acesso em: 17 mai. 2020.

FREITAS JUNIOR, J. C. da S.; MAÇADA, A. C. G.; OLIVEIRA, M.; BRINKHUES, R. A. Big Data e Gestão Do Conhecimento: Definições e Direcionamentos de Pesquisa. **Revista Alcance**, v. 23, n. 4, p. 529-546, 2016. Disponível em: <doi: alcance.v23n4.p529-546>. Acesso em: 08 mar. 2020.

FURLAN, P. K.; LAURINDO, F. J. B. Agrupamentos epistemológicos de artigos publicados sobre big data analytics. **TransInformação**, v. 29, n. 1, p. 91-100, 2017. Disponível em: <<http://www.scielo.br/pdf/tinf/v29n1/0103-3786-tinf-29-01-00091.pdf>>. Acesso em: 08 mar. 2020.

FUSCO, E.; MUCHERONI, M. L. A Inovação nas Organizações Apoiada pela Inteligência Competitiva em Ambiente de Big Data. **Colloquium Socialis** v. 1, n. especial, p.338-343, 2017. Disponível em: <DOI: 10.5747/cs.2017.v01.nesp.s0053>. Acesso em: 02 mai. 2020.

GIMENEZ, M. F. L. **Os Efeitos da Utilização do Big Data na Inteligência Competitiva**. 89 f. Dissertação (Mestrado em Administração) – Universidade Presbiteriana Mackenzie, São Paulo, 2018. Disponível em:

<<http://tede.mackenzie.br/jspui/bitstream/tede/3822/5/MAYCON%20FRANCO%20LOURENÇO%20GIMENEZ.pdf>>. Acesso em: 02 mai. 2020.

GROVER, V.; CHIANG, R. H. L.; LIANG, T-P; ZHANG, D. Creating Strategic Business Value from Big Data Analytics: A Research **Framework**. **Journal of Management Information Systems**, v. 35, n. 2, p. 388-423, 2018. Disponível em: <doi:10.1080/07421222.2018.1451951>. Acesso em: 06 mar. 2020.

HOODA, N.; RANA, P. S.; BAWA, S. Fraudulent Firm Classification: A Case Study of an External Audit. **Applied Artificial Intelligence**, v. 32, n. 1, p. 48-64, 2018. Disponível em: <DOI: 10.1080/08839514.2018.1451032>. Acesso em : 11 jun. 2020.

MANYIKA, J.; CHUI, M.; BROWN, B.; BUGHIN, J.; DOBBS, R.; ROXBURGH, C.; BYERS, A. H. Big data: The next frontier for innovation, competition, and productivity. (Documento online). **McKinsey Global Institute**, 2011. Disponível em: <<https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation>>. Acesso em 07 mar. 2020.

MEIRA, M. F. P. **O impacto da Inteligência Artificial na Auditoria**. 69 f. Dissertação (Mestrado em Contabilidade e Controle de Gestão) – Faculdade de Economia da Universidade do Porto, Portugal, 2019. Disponível em: <<https://repositorio-aberto.up.pt/bitstream/10216/124519/2/368850.pdf>>. Acesso em: 09 mai. 2020.

PEDROSA, I.; LAUREANO, R. M. S.; COSTA, C. J. Motivações dos auditores para o uso das Tecnologias de Informação na sua profissão: aplicação aos Revisores Oficiais de Contas. **RISTI**, n. 15, p. 101-118, 2015. Disponível em: <<http://dx.doi.org/10.17013/risti.15.101-118>>. Acesso em: 26 abr. 2020.

PINTO, H. A. A utilização da inteligência artificial no processo de tomada de decisões: por uma necessária accountability. **Revista de Informação Legislativa**, v. 57, n. 225, p. 43-60, 2020. Disponível em: <https://www12.senado.leg.br/ril/edicoes/57/225/ril_v57_n225_p43>. Acesso em: 07 jun. 2020.

RAJI, I. D.; SMART, A.; WHITE, R. N.; MITCHELL, M.; GEBRU, T.; HUTCHINSON, B.; SMITH-LOUD, J.; THERON, D.; BARNES, P. Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing. **FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency**. Janeiro de 2020, p. 33-44. Disponível em: <<https://doi.org/10.1145/3351095.3372873>>. Acesso em: 11 jan. 2020.

ROTHER, E. T. Editorial: Revisão sistemática X revisão narrativa. **Acta Paulista de Enfermagem**, v. 20, n. 2, 1-2, 2007. Disponível em: <<https://doi.org/10.1590/S0103-21002007000200001>>. Acesso em: 23 mai. 2020.

SAIF, S.; WAZIR, S. Performance Analysis of Big Data and Cloud Computing Techniques: A Survey. **Procedia Computer Science**, v. 132, p. 118-127, 2018. Disponível em: <<https://doi.org/10.1016/j.procs.2018.05.172>>. Acesso em 08 mar. 2020.

SUN, T.; VASARHELYI, M. A. Embracing Textual Data Analytics in Auditing with Deep Learning. *The International Journal of Digital Accounting Research*, v. 18, p. 49-67, 2018. Disponível em: <DOI: 10.4192/1577-8517-v18_3 Accepted March>. Acesso em: 07 jun, 2020.

TAVARES, R. B. Auditoria Interna e *Data Analytics*: o uso da análise inteligente de dados para a detecção de fraudes nas organizações. **Revista Especialize On-line IPOG**, Ano 10, v. , n. 17, 2019. Disponível em: <<https://assets.ipog.edu.br/wp-content/uploads/2019/12/07015620/rosimeire-barbosa-tavares-191201615.pdf>>. Acesso em: 03 mai. 2020.

VEALE, M.; KLEEK, M. V.; BINNS, R. Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making. **Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI'18)**. Abril de 2018. Disponível em: <doi:10.1145/3173574.3174014>. Acesso em: 11 jun. 2020.

VIEIRA, V. F. G. **Auditoria em Tempos de *Big Data & Analytics***: Requisitos Mínimos de Controlo. 95 f. Dissertação (Mestrado em Auditoria Empresarial e Pública) – Instituto Politécnico de Coimbra, Coimbra, 2016. Disponível em: <https://comum.rcaap.pt/bitstream/10400.26/18058/1/Vanessa_Vieira.pdf>. Acesso em: 06 mar. 2020.

VOSGERAU, D. S. R.; ROMANOWSKI, J. P. Estudos de revisão: implicações conceituais e metodológicas. **Revista Diálogo Educacional**, v. 14, n. 41, 165-189, 2014. Disponível em: <<https://periodicos.pucpr.br/index.php/dialogoeducacional/article/view/2317/2233>>. Acesso em: 17 mai. 2020.